Survey of California Vegetation
Classification and Mapping Standards

January 11, 2018

Abstract

The Survey of California Vegetation (SCV) defines the process of vegetation classification and mapping as a series of integrated steps. The process begins with the collection of vegetation data in the field in a standardized format and following established protocols. The field data is then analyzed to produce a hierarchical classification of vegetation types that conforms to the United States National Vegetation Classification Standards (FGDC, 2008). A quantitative, rule-based mapping key is developed and detailed descriptions are written for all vegetation types in the project area. The next step is the creation of a fine-scale digital vegetation map consisting of polygons based on aerial imagery; each polygon is coded with the vegetation type and an array of standardized attributes such as cover, structure and disturbance level. Finally, the accuracy of the completed map is assessed through field surveys and scored following protocols that ensure impartiality. A brief summary of the standards for each of these steps is provided here. For more detail, please follow the links provided within.

Introduction

This document covers requirements for all steps in the integrated vegetation classification and mapping process. It is divided into three sections: Classification, Mapping and Accuracy Assessment. An individual project may be limited to only one step (e.g., creation of a vegetation map based on an existing SCV-compliant classification) and so only one section of this document may apply. A list of the deliverables for a complete vegetation mapping project, including an outline of the report content, is provided in the Vegetation Classification and Mapping Project Deliverables document.

Please direct any questions or comments on these standards to the Department of Fish and Wildlife's (CDFW) Vegetation Classification and Mapping Program (VegCAMP).

Survey of California Vegetation Classification and Mapping Standards
Classification

A. Data Collection for Classification

The SCV process starts with the collection of plant and environmental data in the project area. Spatial information is captured using a GPS device. All other data may be recorded on standard paper forms or an electronic device. Data from the field surveys is transferred to a database, checked for quality and consistency, and prepared for use in the classification and mapping processes.

Sample collection sites are allocated such that multiple samples of the each vegetation type are collected throughout the range of the type within the project area. Sample allocation should employ an analysis that balances three goals: a target number of samples is achieved based on workload predictions for the staff conducting the field surveys; the samples are distributed among the types so that both rare and common types are represented; and access to the collection sites is facilitated based on land ownership and proximity to roads or access trails. VegCAMP has employed several methods for sample allocation. These include Generalized Random Tessellation Stratified (GRTS) survey design (Stevens and Olsen, 2004) and a gradient-directed transect (gradsect) approach as described in Appendix B of the Yosemite National Park Vegetation Classification and Mapping Report (Keeler-Wolf et al., 2012). There is no standard for sample allocation.

In addition to providing the foundation for the hierarchical vegetation classification, the vegetation surveys document plant species locations for use in other applications, such as determining species ranges and abundances.

1. Field Sampling Standards
    a. All vegetation data will be publicly available upon project completion, such as through CDFW's Biogeographic Information and Observation System (BIOS) and other data-sharing utilities; no confidentiality agreements may be made with landowners.
    b. Permission to sample is required from all landowners, public and private. For CDFW-funded projects, written permission is required.
    c. The timing of data collection should coincide with phenology appropriate for the type of vegetation being surveyed, e.g., herbaceous vegetation must be sampled when the herbs are fully developed and flowering/fruiting to obtain accurate cover values and reliable identification.
    d. Generally, for new vegetation types, at least 10 samples should be collected. Field surveys are taken using the "CNPS and CDFW Combined Vegetation Rapid Assessment and Relevé Field Form" and following the "Protocol for Combined Vegetation Rapid Assessment and Relevé Sampling Field Form." Additional information may be collected based on specific project needs, such as evidence of Sudden Oak Death, etc.

e. All data collected in herbaceous vegetation and in newly or provisionally described woody vegetation follows the Relevé protocol. Data collected in previously described shrublands and woodlands follows the Rapid Assessment protocol.
f. Each survey is assigned a unique identifier of eight alpha-numeric digits.
g. Field staff should be trained in vegetation sample design, plant identification, cover value estimation, and how to choose appropriate stands for sampling. They should also be familiar with vegetation classification concepts and with the US National Vegetation Classification Standards (USNVC). Training can be provided by the California Native Plant Society's (CNPS) Vegetation Program.
h. Nomenclature and species codes must conform to the USDA NRCS PLANTS Database (USDA, NRCS, 2015) when possible.
i. All species are identified to the finest taxonomic level possible. All species with 1% cover or more in a plot, or with <1% cover, but occurring in multiple plots, should be identified to genus, species, and infraspecific level when possible.
j. Although not a requirement, voucher specimens should be taken, particularly of plants outside their known range, new invasive species, or other unusual occurrences, and deposited in a recognized herbarium. Care must be exercised when collecting plants that may be rare. See the section on Voucher Collection in "Protocols for Surveying and Evaluating Impacts to Special Status Native Plant Populations and Natural Communities" on the CDFW website for more information.

2. Field Data Entry, Archiving, and Quality Checking Standards
   a. Data from paper field forms is entered into the SCV-standard database framework (an Access database template can be obtained from VegCAMP).
   b. If a digital recording device is used in the field to collect data that would otherwise go on the paper form, the data is structured so that it can be easily imported into the standard database.
   c. Archiving
      (1) Original paper data forms are stored so that they are accessible if any questions should arise concerning the surveys. Scanned copies are also made and archived with other project information.
      (2) Digital ground-level photos taken at each survey location are placed in a folder that is named with the unique identifier of the survey.
   d. Quality control
      (1) The geographic location of each survey is checked to be sure it is in the expected place, i.e., the sample point is on the property listed on the field form and in the general vegetation type recorded.
      (2) The values for all fields are reviewed to be sure they lie within reasonable ranges.
      (3) If data is entered from paper forms, it should meet a 99% accuracy rate for plant species and 95% for other fields, based on a random sample of the data.

B. Data Analysis for Classification

All plant species data collected for a project feed directly into a numerical analysis that clusters similar field samples together based on species composition and dominance. The groups formed and refined during this analysis are the basis of the hierarchical classification that defines the vegetation types used for the project.

1. Data Preparation Standards
   a. All botanical names in the dataset use standard nomenclature and have accompanying plant codes. The USDA NRCS PLANTS Database (USDA, NRCS, 2015) is used as the standard, except in instances where a plant is recognized by The Jepson Manual (second edition), but not NRCS. In such cases, please contact VegCAMP for the proper plant code.
   b. Plant species that appear to be outside their known range are investigated and resolved.
   c. Data entry errors for plant names are checked by reviewing taxa with similar names and/or codes (e.g., ARDO3, *Artemisia douglasii* vs. ARDO4, *Arundo donax*).
   d. Subtle differences in the coding of infraspecific taxa are reviewed to ensure that one taxon was not entered in multiple formats [e.g., *Eriogonum fasciculatum* ssp. *foliolosum* (ERFAF5) versus *Eriogonum fasciculatum* var. *foliolosum* (ERFAF2)].

Note: At this point, the plant species data must be copied into a separate table, so that the detailed plot-related plant data will not be affected.

   e. Taxa identified only to genus or family are reviewed and, in most cases, deleted from the dataset to reduce statistical noise.
   f. Closely related taxa that occur in overlapping ecological settings may be merged into one taxon to reduce statistical noise.
   g. Percent cover estimates for each taxon in the dataset are grouped into predetermined cover classes to normalize the dataset. Braun-Blanquet categories with 7 cover classes are the standard (1=<1%, 2=1-5%, 3=>5-15%, 4=>15-25%, 5=>25-50%, 6=>50-75%, 7=>75%), though occasionally presence-absence categories (0=absent, 1=present) are used.
   h. A table with 3 fields (survey ID, plant code, and cover class) for each field sample is generated.
2. Data Analysis Standards
   (Note: The following steps, a through h, are based on using PC-Ord software. If other software is used, similar program-specific steps should be taken.)
   a. The data from the preceding step is imported into software suitable for analyzing ecological datasets and converted into a plot-by-species matrix.

   b. Statistics that summarize alpha and beta diversity and the coefficient of variation (CV) for both plots and species are run on the plot-by-species matrix. These statistics measure data dispersion or variance in a dataset.
   c. Outlier analyses are run on the plot-by-species matrix to flag outliers that are more than 3 standard deviations from the mean.
   d. Taxa that occur in a small number of plots (i.e., less than 2, 3, 4, and 5 plots) are removed to generate additional plot-by-species matrices. The summary and outlier statistics mentioned in the previous two steps are repeated on these matrices.
   e. A plot-by-species matrix with lower alpha and beta diversity and a lower CV for species (typically <200%) is selected. All species and plot outliers are removed from this matrix.
   f. Agglomerative hierarchical cluster analysis that groups surveys together based on similar frequency and abundance is run on the selected matrix.
   g. Indicator species analysis is run on the cluster diagram to show which species are indicators of each group when the cluster diagram is divided into 2 versus 3 versus 4 versus … 50, etc. groups. Two or three grouping levels that divide the cluster diagram into groups with the lowest average p-values and highest number of significant indicators are selected to drive the vegetation classification process.
   h. For large datasets, especially those with more than 500 samples, it is often recommended that the analysis steps above be run in stages. The initial stage involves analyzing all the samples in the full dataset together and using the cluster output to determine how to break the dataset into more manageable groups or subsets. Additional stages involve repeating the full analysis on each of the smaller subsets.

3. Vegetation Classification

   The vegetation classification is developed based on the cluster groupings from the analyses described above. Ecologists use the frequencies and abundancies of the taxa that hold cluster groupings together ("indicator species") and information from existing classification reports to classify all surveys in the cluster analysis to the lowest level of the hierarchy possible (e.g., Alliance and/or Association). Finer-level cluster groupings guide classification of data to Association, while broader groupings guide classification to Alliance.

   a. After the classification is established, surveys that were deleted from the analysis because they were flagged as outliers are reviewed and, if they share strong similarity with existing types, assigned final classification names.
   b. The vegetation classification is provided to VegCAMP and/or the California Native Plant Society (CNPS) Vegetation Program for review, to ensure compliance with the Manual of California Vegetation (Sawyer et al., 2009) and the USNVC.

C. Vegetation Key, Map Classification, Descriptions, and Crosswalks

Four tools are developed from the hierarchical vegetation classification. The vegetation key is used by mappers to help choose vegetation types for the polygons they draw. The map classification is derived from the vegetation classification, with modifications dependent upon the resolution of the base imagery. Detailed descriptions of all vegetation types found in the project area are written. Crosswalks to other vegetation classification systems are provided so that the results of this project can be compared to other studies.

Examples of a vegetation classification, key, and crosswalks can be found in the "Classification of the Vegetation Alliances and Associations of the Northern Sierra Nevada Foothills, California Vol. 1." Examples of vegetation descriptions are found in these reports: "Classification of the Vegetation Alliances and Associations of the Northern Sierra Nevada Foothills, California Vol. 2" and "Vegetation Classification Manual for Western San Diego County."

1. Vegetation Key
   a. Definitions of terms used in the key, such as absolute cover, dominance, stand, etc., are included at the beginning of the key.
   b. The key is organized by layer (tree, shrub, herb), and then by Group, Alliance, and Association.
2. Map Classification
   a. The map classification is based on the vegetation classification, but is restricted by what is discernable from the aerial imagery. For instance, if the vegetation classification includes an *Atriplex polycarpa* Alliance and an *Atriplex canescens* Alliance, but these two types cannot be distinguished from one another on the aerial imagery, an *Atriplex polycarpa-Atriplex canescens* Mapping Unit may be created in the mapping classification. Since these are non-standard classification units, careful consideration must be given to the naming of mapping units based on their relationships to the nested vegetation classification hierarchy. Ideally, they should nest within the same group or macrogroup.
   b. Each vegetation type may be assigned a map code derived from the order of the types within the map classification.
3. Descriptions
   a. Descriptions of each Alliance and Association should have the following information:
      (1) A summary of the vegetation type, with a short description of the vegetation and the salient environmental characteristics
      (2) The distribution within the study area
      (3) A table of the means and ranges of cover values for each stratum
      (4) A summary of the environmental data
      (5) The number of samples used to describe the type and the sample identifiers
      (6) A discussion of the global distribution of the type, if known

(7) Rarity status (if this is a newly described type, rarity would be estimated in consultation with VegCAMP.)
(8) "Stand tables" summarizing percent constancy and abundance (minimum, maximum and average) values for each species in a type

4. Crosswalks
   a. A crosswalk of the final vegetation types to the US National Vegetation Classification Standard should be provided, organized according to the hierarchical structure of the USNVC.
   b. Crosswalks to other commonly-used classifications, such as California Wildlife Habitat Relations (CWHR), are encouraged but not required.

D. Vegetation Mapping

Vegetation maps consist of geospatially registered polygons which are interpreted through analysis of aerial imagery using head's-up digitizing, image segmentation, or other techniques. Each map polygon is described with the vegetation type along with other attributes, such as cover values and environmental information, which allow the map geodatabase to be analyzed in multiple ways. For example, the conifer tree cover attribute allows one to query the map for all vegetation stands that have conifers, regardless of the species.

Vegetation mappers must have a good ecological knowledge of the vegetation types in the mapping area. This ecological knowledge is reinforced by the field reconnaissance surveys that are performed in order to match vegetation signatures on aerial photographs with actual vegetation types on the ground. The reconnaissance surveys may be augmented by Rapid Assessment and Relevé information compiled during the vegetation classification phase of the project. The vegetation key and descriptions of vegetation types that were developed during the classification phase are additional tools in the mapping process. Ancillary data that can assist the mappers includes datasets such as geology, soils, fire history, topography, or rare plant occurrence maps.

The standards presented here apply to areas being mapped for the first time. For projects that require remapping, please contact VegCAMP to discuss methodology.

1. Mapping Standards
   a. The base imagery for a mapping project should be consistent throughout the project area and ideally should be readily available to all agencies and the public. The base imagery must meet or exceed the NAIP resolution standards (1 meter ground sample distance) and must match the spatial accuracy of the most recent NAIP product. Finer-scale imagery may be used if it can be accessed by the public.
   b. The vegetation type attribution is at least to the Alliance level when possible; if the resolution of the imagery does not allow attribution to that level (as is the case with most herbaceous types), the Group or Macrogroup level is acceptable, depending upon the specific purpose of the map
   c. All vegetation cover is reported as absolute percent cover, based on bird's-eye view (what an interpreter can see from the sky looking down).
   d. Cover values for woody vegetation are recorded in 1% increments.
   e. Cover values for herbaceous vegetation are recorded in the following cover classes: <2%, 2-9%, 10-40%, >40%.
   f. The minimum mapping unit (MMU) varies with the size of the project, but is not greater than 10 acres and is usually 1 or 2 acres; wetlands and other special types are mapped at ¼ acre. The minimum width of polygons depends on the MMU, but is generally no less than 30 feet (10 meters).

g. Rules for aggregating stands that are below MMU
   (1) A below-MMU vegetation unit that is completely surrounded by another vegetation type is aggregated with the surrounding type.
   (2) Similar growth forms are combined when possible: tree-dominated types are aggregated with other tree-dominated types, shrub types with other shrub types, and herbaceous types with other herbaceous types.
   (3) If a below-MMU vegetation unit is the same growth form as two adjacent larger stands, and the adjacent stand types are very dissimilar in environment, the unit may be aggregated with the more environmentally similar type.
   (4) Whenever possible, wetland vegetation types are not aggregated with upland types, even if they are in the same growth form.

h. Rules for dividing polygons
   A polygon of a single vegetation map unit should be divided into smaller polygons based upon a change in cover class. Even though cover is attributed in 1% increments, the recommended cover classes to be used for dividing polygons of woody vegetation are the Braun-Blanquet categories (1=<1%, 2=1-5%, 3=>5-15%, 4=>15-25%, 5=>25-50%, 6=>50-75%, 7=>75%). The cover classes used to determine divisions of herbaceous vegetation are: <2%, 2-9%, 10-40%, >40%.
   (1) Overstory cover break
      (a) Break a polygon on overstory cover if there is a change in cover class of the dominant/nominal layer. For example, if the vegetation is shrub-dominated, a change in cover class of the shrub layer is reason for a polygon division.
      (b) For projects with a 1-acre MMU, there is typically a 3-acre MMU for a break in the overstory cover, i.e., the resulting polygons must be at least 3 acres in size.
   (2) Understory cover break
      (a) Break a polygon on understory cover if there is a change in cover class of the understory layer. For example, if the vegetation is tree-dominated, a change in cover class of the shrub layer or a cover class break in the herbaceous layer as in (h.) above is reason for a polygon division.
      (b) For projects with a 1-acre MMU, there is typically a 5-acre MMU for a cover class break in the understory.
   (3) Height and size class break for tree types
      (a) Break a polygon on tree height if there is a change in the height class or CWHR size class of the tree layer.
      (b) For projects with a 1-acre MMU, there is typically a 3-acre MMU for a break in tree height or size class.

i. Required mapping attributes include:
   (1) Vegetation type
   (2) Conifer cover

(3) Hardwood cover
(4) Total tree cover
(5) Total shrub cover. Depending on the imagery used this may not be interpretable over a certain percent cover of overstory trees.
(6) Total herb cover. Depending on the imagery used this may not be interpretable over a certain percent cover of overstory trees or shrubs.
(7) Non-native species cover (see mapping template geodatabase)
(8) Roadedness
(9) Clearing
(10) Attribution method: the polygon was attributed through image interpretation, based on a field survey, or based on a less-formal field reconnaissance
(11) CWHR size class for trees (can be modeled based on field data)
(12) Tree height (can be modeled based on field data)
(13) Other project-specific attributes (such as other disturbances, sudden oak death, etc.)

2. Geodatabase Quality Checking
   a. A topology check is done to eliminate gaps and overlaps.
   b. Acreage for all polygons is checked to identify slivers and polygons that are below the MMU. Polygons that are smaller than the MMU are reviewed to determine if they should be merged with adjacent polygons or kept as exceptions (e.g., polygons that are truncated by the project boundary).
   c. Attributes are checked to make sure there is a value for every attribute. “N/A” rather than a NULL value is used if an attribute does not apply to a specific polygon.
   d. Attributes are checked for obvious errors such as nonsensical cover values (e.g. is the total tree cover lower than the conifer cover and hardwood cover combined? Is there a shrub cover value for all shrub type polygons?).
   e. If a notes field is used, it is checked to ensure that all information is relevant and useful. Any questions or notes that were used during the mapping process are removed. Abbreviations are changed to fully spell out the intended words.
3. Geodatabase Metadata
   a. The following should be included in the vegetation map metadata:
      (1) A summary describing the purpose of the mapping project, how the project boundary was determined, the mapping classification used and level of the hierarchy mapped.
      (2) The date and source of the base imagery and the field samples used in the mapping project. Also list imagery or other datasets used as ancillary information.
      (3) Contact information for an individual or organization that is knowledgeable about the dataset, including the name, telephone number and/or email.

(4) A full description of each field and its potential values. Note whether the value was directly interpreted or modeled, and if modeled, how it was modeled. If an attribute is not required for a given vegetation or land use type, note that it may have the value “not applicable.”
(5) Minimum mapping unit for all mapping types and minimum polygon width, including MMUs for overstory and understory cover breaks and other attributes.
(6) How polygons on the mapping area’s boundary were mapped (e.g., only up to the edge, with a buffer, or so the entire stand is mapped), attributed (e.g., is disturbance for the entire stand included, or just evaluated within the portion of the stand within the project boundary), and edge-mapped or seamed with an existing map.
(7) Details of how the map was seamed or edgematched to adjacent vegetation maps.
(8) Any access or use constraints on the map.
(9) A description of the Accuracy Assessment method and results.
(10) Citations and web addresses for project reports
(11) Update frequency, if the map will be updated
(12) The physical and/or digital location of any field datasets such as sample point locations, Relevé/Rapid Assessment field data forms, field photos, or herbarium specimens

E. Map Accuracy Assessment

An Accuracy Assessment (AA) analysis helps the map users determine how much confidence can be assigned to each of the mapped vegetation types, and provides an understanding of the map's appropriateness for various applications.

Polygons are selected for assessment from the draft vegetation map and are presented to the field crew for survey. In order to prevent bias during the field surveys, the AA field map shows only the outlines of the polygons to be visited; vegetation type and map attributes are not shown. The field observations are then compared to the draft map attributes and each polygon is scored. To ensure the impartiality of the Accuracy Assessment process, members of the field survey team cannot have mapped the area they survey and scorers cannot have mapped or surveyed the area they score.

The mappers are given the results of the Accuracy Assessment, including scores and comments for specific polygons, so that both specific and systematic errors can be corrected prior to map finalization.

An example of the Accuracy Assessment process, including goodness-of-fit (also known as fuzzy logic) scoring, can be found in "Northern Sierra Nevada Foothills Project: Vegetation Mapping Report."

1. Accuracy Assessment Standards
   a. The Survey of California Vegetation requires 80% overall accuracy for vegetation maps.
   b. Every map must be verified. However, the formal Accuracy Assessment can be waived if at least 40% of all of mapped polygons have been field surveyed, either before mapping or after completion of the draft map.
2. Sample Allocation
   Sample allocation should employ an analysis that balances four goals:
   a. Achieving a target number of samples based on workload predictions for the staff conducting the field surveys
   b. Distributing the samples amongst the vegetated mapping types so that both rare and common types are represented
   c. Facilitating access to vegetation polygons based on land ownership and proximity to roads or access trails
   d. Ensuring that previously surveyed polygons are not revisited
3. Accuracy Assessment Field Surveys
   VegCAMP does not have a standard field form for Accuracy Assessments, however, examples of two different formats for collecting data can be found in the appendices of these reports: "Northern Sierra Nevada Foothills Project: Vegetation Mapping Report" and "2013 California Desert Vegetation Map and Accuracy Assessment in Support of the Desert Renewable Energy Conservation Plan."

   a. For a polygon to be assessed for accuracy, at least 10-20% of its area should be viewed. If this is not possible, the vegetation type and other attributes at the survey point should be recorded as information to the mapper. If less than 100% of the polygon can be viewed, field crews should note whether environmental features such as aspect, slope, and overall topography are consistent throughout the polygon, implying that the entire polygon is likely to be the same vegetation type.
   b. The vegetation type is determined using the key and descriptions developed in section C: Vegetation Key, Map Classification, Descriptions, and Crosswalks. A list of dominant species and their corresponding percent cover estimates are recorded, along with values for other mapped attributes.
   c. The spatial accuracy of the polygon is assessed and instructions are provided on how to redraw the lines if adjustments are necessary.
   d. If the polygon contains more than one vegetation type, notes are provided as to how the polygon should be divided and a separate survey is taken for each type.
   e. AA field data is entered into a database to facilitate scoring (an MS Access database template can be obtained from VegCAMP).

4. Scoring
   a. Each field-assessed polygon is scored by comparing the vegetation type determined in the field survey to the type assigned by the mapper. A score of 0-5 is given based on a “goodness-of-fit” approach, where ecological and floristic similarities are considered.
   b. Other information gathered in the field is provided to the mappers for reference, but does not contribute to the score.
   c. If 80% accuracy is not attained for a particular vegetation type there are several options:
      (1) The mapper can decide to aggregate to a less specific hierarchical unit (e.g., from Association to Alliance level; from Alliance to Group level).
      (2) The mapper can choose to keep the vegetation type, but report that its accuracy is lower than typically acceptable.
      (3) Mappers can make specific and systematic corrections to the polygon attributes or delineations based on the AA results.

Literature Cited

Federal Geographic Data Committee (FGDC). 2008. National Vegetation Classification Standard, Version 2 FGDC-STD-005-2008. Vegetation Subcommittee, FGDC Secretariat, U.S. Geological Survey. Reston, VA. 55 pp. + Appendices.

Keeler-Wolf, Todd, Peggy E. Moore, Edward T. Reyes, John M. Menke, Deborah N. Johnson and Denise L. Karavidas. 2012. Yosemite National Park Vegetation Classification and Mapping Report. Natural Resource Technical Report NPS/YOSE/NRTR—2012/598. US Department of the Interior National Park Service, Fort Collins, CO.

Sawyer, J.O., T. Keeler-Wolf, and J.M. Evens. 2009. A Manual of California Vegetation, Second Edition. California Native Plant Society, Sacramento. 1300 pp.

Stevens, Don L. and Anthony R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99:262-278.

USDA, NRCS. 2015. The PLANTS Database (http://plants.usda.gov, 4 June 2015). National Plant Data Team, Greensboro, NC 27401-4901 USA.